DOCUMENT RESUME

ED 089 544 FL 004 873

AUTHOR Paper, Herbert H.; Jazayery, Mohammad Ali
TITLE The Writing System of Modern Persian. American Council of Learned Societies Program in Oriental Languages, Publication Series B - Aids - Number 4.
INSTITUTION American Council of Learned Societies, New York, N.Y.; Spoken Language Services, Inc., Ithaca, N.Y.
SPONS AGENCY Ford Foundation, New York, N.Y.
PUB DATE 55
NOTE 36p.
AVAILABLE FROM Spoken Language Services, Inc., P. O. Box 783, Ithaca, N. Y. 14850 ($3.00)

EDRS PRICE MF-$0.75 HC Not Available from EDRS. PLUS POSTAGE
DESCRIPTORS Language Instruction; Letters (Alphabet); *Manuscript Writing (Handlettering); Modern Languages; *Morphology (Languages); *Persian; Printing; Punctuation; *Sentence Structure; Student Writing Models; Textbooks; Uncommonly Taught Languages; *Written Language

ABSTRACT

This manual describes and illustrates the writing system of modern Persian. It is modelled after Frank A. Rice's "Classical Arabic: The Writing System." Five sections concerning the writing system are provided: (1) printed Persian, (2) typewritten Persian, (3) handwritten Persian, (4) the combination of words, and (5) punctuation. (RL)

ED 089544

THE WRITING SYSTEM OF MODERN PERSIAN

Herbert H. Paper
and
Mohammad Ali Jazayery

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Spoken Language Services, Inc.

FL 004 873

Copyright, 1955, by

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY
Spoken Language Services, Inc.
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER"

AMERICAN COUNCIL OF LEARNED SOCIETIES
Program in Oriental Languages
Publication Series B — Aids — Number 4

THE WRITING SYSTEM OF MODERN PERSIAN

Herbert H. Paper
and
Mohammad Ali Jazayery

American Council of Learned Societies
Washington, D.C.
1955

The research and compilation of which this work is a result were brought to completion under a subvention from the Board on Overseas Training and Research (Ford Foundation)

LITHOGRAPHED IN THE UNITED STATES OF AMERICA

PREFACE

This manual, originally prepared by Professor Herbert H. Paper, of the University of Michigan, is closely modelled on Frank A. Rice's Classical Arabic: The Writing System (Foreign Service Institute, 1952). It has been revised to some extent by the following persons: Charles A. Ferguson, Carleton T. Hodge, Mohammad Meimandi-Nejad, and Mohammad Ali Jazayery. The notes on typewritten Persian, handwritten Persian, the combination of words, and punctuation are Mr. Jazayery's.

TABLE OF CONTENTS

			Pages
PRINTED PERSIAN			1-19
0	General	1	
1-2	The Alphabet	1	
3	Vowels	3	
4	Non-Connectors	4	
5-16	Connectors	5	
17	Ligatures	16	
18-26	Special Signs	16	
27	Numbers	19	
TYPEWRITTEN PERSIAN			20-21
HANDWRITTEN PERSIAN			22-24
COMBINATION OF WORDS			25-26
PUNCTUATION			27-30

PRINTED PERSIAN

GENERAL

0. Persian is customarily written and printed in an adapted form of the Arabic alphabet consisting of 32 letters and a number of signs. It differs from its Arabic original in that it has added four new letters to take care of Persian sounds which do not exist in Arabic. In general, all words of Arabic origin, as well as a number of Arabic phrases common in Persian, appear in their original Arabic orthography.

While the letters of the alphabet are basically the same in form whether written, printed, or typed, there are slight differences. For our purposes, three main types may be distinguished: printed, typed, and handwritten (which is sometimes, though not so frequently as in earlier times, used in lithography or photo-printing). In the present treatment we are mainly concerned with the printed variety of the Persian alphabet. However, a short note on handwritten Persian has been added to serve as an introductory guide to the reading of this style. The typewritten form is also briefly discussed.

1. Persian is written and printed from right to left.
2. All of the consonants, but only three of the six simple vowels, are indicated.
3. The manner of writing and printing is cursive, and most of the letters are connected to preceding and following letters within the same word.
4. There are no capital letters.
5. Most of the letters have four variant shapes, depending on whether or not they are connected to preceding and/or following letters. Later on the four shapes will be called initial, medial, final, and independent.
6. Some letters are identical in basic form and are distinguished from one another by the arrangement of one, two, or three dots placed above, below, or inside the basic form.

THE ALPHABET

1. The following list shows the letters of the Persian alphabet in the shape they have when they stand alone (= independent), that is, not connected to a preceding or following letter. This is the conventional order of the alphabet according to which the dictionaries are arranged.

The Persian Alphabet

Letter	Name	Transcribed by
ا	ʾælef	ʾ, a (or no symbol)
ب	be	b
پ	pe	p
ت	te	t
ث	se	s
ج	jim	j
چ	ce	c
ح	he (hottí)	h
خ	xe	x
د	dal	d
ذ	zal	z
ر	re	r
ز	ze	z
ژ	že	ž
س	sin	s
ش	šin	š
ص	sad	s
ض	zad	z
ط	ta	t
ظ	za	z
ع	ʾeyn	ʾ (or no symbol)
غ	qeyn	q
ف	fe	f
ق	qaf	q
ک	kaf	k
گ	gaf	g
ل	lam	l
م	mim	m
ن	nun	n
و	vav	v, u, o, ow
ه	he (hævvéz)	h, e
ی	ye	y, i

Notice that in a number of instances several letters represent the same sound in Persian. In these cases, only one of these letters ordinarily occurs in words of Persian origin, while all of them can be found in Arabic loan words. However, in a few cases, even words which are not Arabic borrowings are spelled with letters otherwise peculiar to Arabic. A number of Persian proper nouns (both place names and personal names) are spelled with typical Arabic letters by some people and with the ordinary letters by others. The present tendency is in favor of the latter type of spelling. The same situation exists in the case of words borrowed from languages other than Arabic (especially Greek and Turkish).

Persian Sound	Letters		
z	ز	ze	
	ذ	zal	Chiefly in Arabic loan words
	ض	zad	
	ظ	za	
s	س	sin	
	ث	se	Chiefly in Arabic loan words
	ص	sad	
t	ت	te	
	ط	ta	Chiefly in Arabic loan words
h	ح	he hottí	Chiefly in Arabic loan words
	ه	he hævvéz	
q	غ	qeyn	
	ق	qaf	Chiefly in Arabic loan words
'	ا	'æléf	
	ع	'eyn	Chiefly in Arabic loan words

2. The following shows the Persian letters in their alphabetical order and written from right to left. Note the placing of the letters relative to the line of writing.

ا ب پ ت ث ج چ ح خ د ذ ر ز ژ

س ش ص ض ط ظ ع غ ف ق ك گ ل

م ن و ه ی

VOWELS

3. Persian has six vowels: a, æ, e, i, o, u. Two of them, i and u, are regularly indicated in the writing by the letters ی (ye) and و (vav) re-

spectively. One, a, is regularly indicated by ا ('æléf) in the middle or at the end of a word. At the beginning it is sometimes indicated by a special mark ◌̄ over ا ('æléf): آ. The other three vowels are not usually represented in the writing. Four special signs, described in Section 18, may be used to indicate æ, e, and o, or absence of vowel, but these signs are rarely used in ordinary written or printed Persian. Occasionally o is indicated by و (vav); and at the end of a word e is sometimes indicated by ه (he hævváz).

There is also a special sign to indicate double consonants (see Section 23). This sign is not ordinarily used.

NON-CONNECTORS

4. The following letters are non-connectors. They do not have a cursive linking to a following letter, though they do connect to a preceding letter if the preceding letter is a connector.

ا 'æléf is characteristically a vertical stroke.

د dal

ذ zal is exactly like dal, but with one dot above it.

ر re

ز ze is exactly like re, but with one dot above it.

ژ že is exactly like re, but with a dot triangle above it.

و vav

At the beginning of a word, the letter ا ('æléf) indicates '. A vowel naturally follows, as Persian has no words beginning with two consonants. 'u is represented by او, 'i by ای. 'a has a special symbol over the ا ('æléf): آ (see Section 19.10). 'æ, 'o, and 'e are ordinarily written with ا ('æléf) alone.

In the middle of a word, ا ('æléf) most often represents the vowel a, but it may also represent '. At the end of a word, ا ('æléf) nearly always indicates the vowel a (see Section 24).

The symbol ◌̄ (Section 19.11) for 'a is also used in the middle and at the end of words.

The letter و (vav) at the beginning of a word represents the consonant v; elsewhere it may represent u, o, ow, or v.

از 'æz 'from'	در dær 'in'	داد dad 'he gave'
او 'u 'he'	دو do 'two'	درد dærd 'pain'
ارزش 'ærzéš 'value'	دز dez 'fortress'	دوا dævá 'medicine'
		دود dud 'smoke'

اردو 'ordú 'Urdu'	روز ruz 'day'	ورم værém 'swelling'
ادوار 'ədvár 'periods'	زد zæd 'he struck'	ورود vorúd 'arrival'
	زود zud 'quick'	
	زور zur 'strength'	
	زرد zærd 'yellow'	

ژ dež 'fortress'

ژوژ ruž 'rouge'

CONNECTORS

5. All other letters connect both to preceding letters (except when they are non-connectors) and to following letters. These may be arranged according to their basic forms.

6. b, n, y, t, p, s. These letters are all characterized by an upward stroke with one to three dots above or below. The medial forms are as follows:

ـبـ b	ـیـ y	ـپـ p
ـنـ n	ـتـ t	ـسـ s

In their final and independent forms, there is significant variation only in the basic forms of ن n and ی y.

6.10 b (be) has a single dot below.

6.11 b- has the initial shape: بـ

باز	baz 'open'	ابرو	'æbrú 'eyebrow'
بد	bæd 'bad'	ابر	'æbr 'cloud'

6.12 -b- has the medial shape: ـبـ

ببر	bæbr 'tiger'	تبریز	tæbríz 'Tabriz'
تبر	tæbér 'axe'	زیبا	zibá 'beautiful'

6.13 -b has the final shape: ـب

زیب	zib 'ornament'	ادیب	'ədíb 'man of letters'
تب	tæb 'fever'	اریب	'oríb 'slanting'

6.14 b has the independent shape: ب

تاب	tab 'glow'	باب	bab 'gate'
ناب	nab 'pure'	ترب	torób 'radish'

6.20 n (nun) has one dot above.

6.21 n- has the initial shape: نـ

نر	nær 'male'	نیز	niz 'also'
ناو	nav 'ship'	نو	now 'new'

6.22 -n- has the medial shape: ـنـ

بنیاد	bonyád 'foundation'	هنوز	hænúz 'yet, still'
پنیر	pænír 'cheese'	تند	tond 'fast'

6.23 -n has the final shape: ن

تن	tæn 'body'	دین	din 'religion'
بن	bon 'root'	زودترین	zudtærín 'quickest'

6.24 n has the independent shape: ن

نان	nan 'bread'	زدن	zædán 'to strike'
نادان	nadán 'ignorant'	دندان	dændán 'tooth'

6.25 Written medial or final nb is pronounced mb:

زنبور	zæmbúr 'wasp'	انبر	'æmbór 'pincers'

6.30 y (ye) has two dots below (when initial or medial).

The letter ی (ye) at the beginning of a word represents the consonant y; elsewhere it may represent i or y.

The sequence ای ('æléf-ye) at the beginning of a word often represents 'i, and sometimes it represents 'ey.

6.31 y- has the initial shape: ـیـ

یا	ya 'or'	یونان	yunán 'Greece'
ایران	'iran 'Iran'	این	'in 'this'

6.32 -y- has the medial shape: ـیـ

نیز	niz 'also'	پنیر	pænír 'cheese'
بیابان	biabán 'desert'	بین	beyn 'between'

6.33 Final ye is deeper than the final forms of the other letters treated in Section 6 (except n). It is also characterized by the curving connecting line. Notice that the two dots are ordinarily not written below the final and independent forms

ایرانی	'iranı 'Iranian'	یونانی	yunanı 'Greek'
بی	bi 'without'	نادانی	nadanı 'ignorance'

6.34 y has the independent shape: ی

روزی	rúzi 'a day'	زردی	zærdí 'yellowness'
بازی	bazí 'play, game'	یاری	yarí 'help'

6.35 At the end of certain Arabic loan words the letter ی represents the sound a. You will have to learn the words that are written with ی for a instead of ا ('æléf). (See Section 26).

6.40 t (te) has two dots above.

6.41 t- has the initial shape: ـتـ

تابوت tabút 'coffin' — تب tæb 'fever'

تبر tæbér 'axe' — تبریز tæbríz 'Tabriz'

6.42 -t- has the medial shape: ـتـ

پتو pætú 'blanket' — متر metr 'meter'

دکتر doktór 'doctor' — کمتر kæmtér 'less'

6.43 -t has the final shape: ـت

تربیت tærbiyét 'education' — بت bot 'idol'

پوست pust 'skin' — بیست bist 'twenty'

6.44 t has the independent shape: ت

تابوت tabút 'coffin' — عبارت 'ebarét 'phrase'

احساسات 'ehsasát 'feelings' — لغات loqát 'words'

6.50 p (pe) has three dots arranged in a triangle below.

6.51 p- has the initial shape: ـپـ

پا pa 'foot' — پارو parú 'oar'

پر por 'full' — پر pær 'feather'

6.52 -p- has the medial shape: ـپـ

تپید tæpíd 'palpitated' — بپرد. bépæræd 'Let him fly.'

بپر. bépær 'Fly!' — بپز. bepéz 'Cook!'

6.53 -p has the final shape: ـپ

پیپ pip '(tobacco) pipe' — چپ cæp 'left'

تیپ tip 'brigade' — جیپ jip 'jeep'

6.54 p has the independent shape: پ

پاپ pap 'Pope' — چاپ cap 'printing'

توپ tup 'cannon' — کرپ kerép 'crepe'

6.60 s (se) has three dots above.

6.61 s- has the initial shape: ـسـ

ثابت sabét 'proved; fixed' — ثبت sæbt 'registration'

ثروت særvét 'wealth' — ثلث sols 'a third'

6.62 -s- has the medial shape: ـسـ

نثر næsr 'prose' — تثنیه tæsniyé 'dual (gram.)'

مثال mesál 'example' — مثبت mosbét 'affirmative'

6.63 -s has the final shape: ث

عبث 'abés 'futile' ثالث salés 'third'

مثلث mosællés 'triangular' تثليث tæslís 'trinity'

6.64 s has the independent shape: ث

وارث varés 'heir' توارث tævarós 'inheritance'

ورّاث vorrás 'heirs' ارث 'ers 'heritage'

7. j, c, x, h. These letters, like those of Section 6, have the same basic form but differ only in the arrangement of dots.

ـجـ j ـچـ c ـحـ h ـخـ x

7.10 j (jim) is characterized by a dot placed below the basic form.

7.11 j- has the initial shape: جـ

جا ja 'place' جديد jædíd 'new'

جز joz 'except' تاجر tajér 'merchant'

7.12 -j- has the medial shape: ـجـ

بيجان biján 'lifeless' زنجير zænjír 'chain'

رنجيدن rænjidén 'to be offended' اينجا 'injá 'here'

7.13 -j has the final shape: ـج

پنج pænj 'five' رنج rænj 'suffering'

برنج berénj 'rice' خليج xælíj 'gulf'

7.14 j has the independent shape: ج

تاج taj 'crown' زاج zaj 'vitriol'

برج borj 'tower' خرج xærj 'expense'

7.20 x (xe) is characterized by a dot placed above the basic form.

7.21 x- has the initial shape: خـ

خار xar 'thorn' خان xan 'khan'

خر xær 'donkey' خنجر xænjér 'dagger'

7.22 -x- has the medial shape: ـخـ

بخت bæxt 'luck' بخاری boxarí 'stove'

پختن poxtén 'to cook' نخود noxód 'peas'

7.23 -x has the final shape: ـخ

بيخ bix 'root' نخ næx 'thread'

يخ yæx 'ice' زنخ zænéx 'chin' (literary)

7.24 x has the independent shape: خ

نرخ nerx 'rate' رخ rox 'face'

چرخ cærx 'wheel' کاخ kax 'palace'

7.25 The sequence xa is in certain words written خوا . The و (vav) here is an historical writing and has no sound value.

خواب xab 'sleep' خوابیدن xabidén 'to sleep'

خواندن xandén 'to read' خواهر xahér 'sister'

7.26 Note that the letter و (vav) after a خ (xe) sometimes represents the vowel o.

خود xod 'self' خورشید xoršíd 'the sun'

نخود noxód 'peas' خورد xord 'he ate'

7.30 c (ce) is characterized by three dots placed below the basic form.

7.31 c- has the initial shape: چ

چاپ cap 'print(ing)' چای cay 'tea'

چپ cæp 'left' چرا céra 'why'

7.32 -c- has the medial shape: ـچـ

پیچیدن picidén 'to twist' بچرخ bécærx 'Turn!'

7.33 -c has the final shape: ـچ

پیچ pic 'twist, screw' هیچ hic 'nothing'

7.34 c has the independent shape: چ

پوچ puc 'vain, futile' کوچ kuc 'migration'

7.40 h (he hotti) has the same shape in all positions as j, c, and x except that it has no dots of any kind and is simply the basic form of this series.

7.41 h- has the initial shape: ه

حجاز hejáz 'Hejaz' حزب hezb 'political party'

حیوان heyván 'animal' احتیاج 'ehtiyáj 'need'

7.42 -h- has the medial shape: ـهـ

بحر bæhr 'sea' تحت tæht 'underneath'

تحریر tæhrír 'writing' ریحان reyhán 'sweet basil'

7.43 -h has the final shape: ـه

ذبح zebh 'act of slaughtering' شبح šæbéh 'phantom'

7.44 h has the independent shape: ـه

نوح nuh 'Noah' شرح šærh 'explanation'

8. s and š (sin and šin) have the same basic form. The sin is dotless, and the šin has a triangle of dots above.

8.10 s- has the initial shape: س

ساختن saxtén 'to build'　　اسب 'æsb 'horse'

استاد 'ostád 'professor'　　استخر 'estæxr 'pool'

8.11 -s- has the medial shape: ـس

نیست nist 'is not'　　جستن jostén 'to search'

بیست bist 'twenty'　　بستن bæstén 'to tie'

8.12 -s has the final shape: س

بس bæs 'enough'　　سپس sepæs 'then'

جنس jens 'species; sex'　　پس pæs 'then'

8.13 s has the independent shape: س

اساس 'æsás 'basis'　　ترس træs 'fear'

خروس xorús 'rooster'　　یاس yas 'jasmine'

8.20 š- has the initial shape: ش

شاید šáyæd 'perhaps'　　شب šæb 'evening; night'

شدن šodén 'to become'　　شستن šostén 'to wash'

8.21 -š- has the medial shape: ـش

پشت pošt 'back'　　جشن jæšn 'festival'

خشنود xošnúd 'pleased'　　نشستن nešæstén 'to sit'

8.22 -š has the final shape: ش

خویش xiš 'self'　　خیش xiš 'plough'

ریش riš 'beard'　　زنش zæneš 'his wife'

8.23 š has the independent shape: ش

خوش xoš 'happy'　　دوش duš 'shower'

ارزش 'ærzéš 'value'　　ورزش værzéš 'sport'

9. k and g (kaf and gaf) have the same basic form. Kaf has one oblique stroke in its initial and medial forms.

9.10 k- has the initial shape: ک

کتاب ketáb 'book'　　کار kar 'work'

کارد kard 'knife'　　دکتر doktór 'doctor'

9.11 -k- has the medial shape: ـک

شکار šekár 'hunting'　　شکر šekær 'sugar'

تکرار tekrár 'repetition'　　تشکر tæšekkór 'thanks'

9.12 -k has the final shape: ك

یك yek 'one'　　خنك xonók 'cool'

اشك 'æšk 'tear'　　بانك bank 'bank'

9.13 k has the independent shape: ك

پاك pak 'pure'	ترك tork 'Turk'
خاك xak 'soil; dust'	تریاك teryák 'opium'

9.20 gaf differs from kaf only in having two parallel oblique strokes in all its forms.

9.21 g- has the initial shape: گ

گرگ gorg 'wolf'	گیج gij 'dizzy'
گذشتن gozæštén 'to pass'	گاو gav 'cow'

9.22 -g- has the medial shape: ـگـ

جگر jegér 'liver'	تگرگ tægærg 'hail'
تنگی tængí 'narrowness'	سنگین sængín 'heavy'

9.23 -g has the final shape: ـگ

ریگ rig 'sand'	زنگ zæng 'bell'
جنگ jæng 'war'	سگ sæg 'dog'

9.24 g has the independent shape: گ

گرگ gorg 'wolf'	بزرگ bozórg 'big'
تگرگ tægærg 'hail'	برگ bærg 'leaf'

10. f and q (fe and qaf) differ only in the number of dots and in their final and independent forms, where qaf has a deeper loop than fe; the final form of fe is hollow, while that of qaf is not.

10.10 f- has the initial shape: ف

فرزند færzænd 'child'	فکر fekr 'thought'
افتادن 'oftadæn 'to fall'	افسر 'æfsær 'officer'

10.11 -f- has the medial shape: ـفـ

سفید sæfíd 'white'	نفت næft 'oil'
جفت joft 'pair'	تفنگ tofæng 'rifle'

10.12 -f has the final shape: ـف

شریف šæríf 'noble'	کثیف kæsíf 'dirty'
کشف kæšf 'discovery'	کیف kif 'purse, bag'

10.13 f has the independent shape: ف

برف bærf 'snow'	شگرف šegærf 'wonderful'
حرف hærf 'letter of the alphabet'	ناف naf 'navel'

10.20 qaf differs from fe in having two dots placed above instead of one.

10.21 q- has the initial shape: ق

قانون qanún 'law'
قبر qæbr 'grave'
قدری qædri 'some'
قرن qærn 'century'

10.22 -q- has the medial shape: ـقـ

نقاش næqqáš 'painter'
سقا sæqqá 'water-carrier'
نقد næqd 'cash'
منقار menqár 'beak'

10.23 -q has the final shape: ـق

دقیق dæqíq 'minute (adj.); careful'
رفیق ræfíq 'friend'
زنبق zæmbéq 'iris'
تفریق tæfríq 'subtraction'

10.24 q has the independent shape: ق

برق bærq 'electricity, lightning'
شرق šærq 'east'
قاچاق qacáq 'contraband; smuggle'
اتفاق 'ettefáq 'incident'

11. l (lam) is characteristically a vertical stroke. It differs from ا ('æléf) chiefly by connecting to a following letter.

11.10 l- has the initial shape: لـ

لندن lændén 'London'
لگد lægéd 'kick (noun)'
لخت loxt 'naked'
لباس lebás 'clothes'

11.11 -l- has the medial shape: ـلـ

جلد jeld 'skin, hide'
تلخ tælx 'bitter'
جلو jelów 'front'
گلو gælú 'throat'

11.12 -l has the final shape: ـل

گل gol 'rose'
نیل nil 'Nile'
پل pol 'bridge'
بلبل bolból 'nightingale'

11.13 l has the independent shape: ل

دل del 'heart'
پول pul 'money'
دول dovél 'governments'
دنبال dombál 'rear'

12. m (mim) is characteristically a small circle, often merely a dot on a downward break in the line.

12.10 m- has the initial shape: مـ

ما ma 'we'
مادر madér 'mother'
مرد mærd 'man'
میز miz 'table'

12.11 -m- has the medial shape: ـمـ

محمد mohæmméd 'Mohammed'
بیمار bimár 'sick'
سمت sæmt 'direction'
شیمی šimí 'chemistry'

12.12 -m has the final shape: م

نیم nim 'half'
سیم sim 'wire'
کتابم ketábæm 'my book'
رفتم. ræftæm 'I went.'

12.13 m has the independent shape; م

تمام tæmám 'whole'
سلام sælám 'hello'
سوم sevvóm 'third'
مردم mærdóm 'people'

13. h (he hævvéz) does not have any one characteristic form.

13.10 h- has the initial shape: ه

هر hær 'each, every'
هفت hæft 'seven'
هشت hæšt 'eight'
دهان dæhán 'mouth'

13.11 -h- has the medial shape: ـهـ

بهتر behtér 'better'
بهار bæhár 'spring'
بهشت behéšt 'paradise'
چهل cehél 'forty'

13.12 -h has the final shape: ـه. The final h (ه) often indicates the vowel e.

سه se 'three'
رفته ræfté 'gone'
تربچه torobcé 'radish'
تخته tæxté 'board, plank'

13.13 h has the independent shape: ه. The independent h (ه), too, often indicates the vowel e.

راه ráh 'path, road'
اجازه 'ejazé 'permission'
پیاده piadé 'on foot'
ده deh 'village'

14. s and z (sad and zad) have the same basic form. The sad is dotless; the zad has one dot above.

14.10 s has the initial shape: ص

صابون sabún 'soap'
صبح sobh 'morning'
صحیح sæhíh 'correct'
صد sæd 'hundred'

14.11 -s- has the medial shape: ـصـ

تصمیم tæsmím 'decision'
بصره bæsré 'Basra'
مصر mesr 'Egypt'
سیصد sisæ[illegible]ree hundred'

14.12 -s has the final shape: ـص

نقص næqs 'defect'
ناقص naqés 'defective'
شخص šæxs 'person'
مرخص moræxxæs 'permitted to leave'

14.13 s has the independent shape: ص

مخصوص mæxsús 'special'　　خلوص xolús 'sincerity'

رقاص ræqqás 'dancer'　　خاص xass 'particular'

14.20 z- has the initial shape: ض

ضرر zærær 'loss'　　ضيافت ziyafét 'banquet'

اضافه 'ezafé 'addition'　　واضح vazéh 'clear, plain'

14.21 -z- has the medial shape: ـض

هضم hæzm 'digestion'　　اقضا 'enqezá 'expiration'

قضاوت qezavét 'judgment'　　مضر mozér 'harmful'

14.22 -z has the final shape: ـض

نقض næqz 'violation'　　محض mæhze 'only for'

نبض næbz 'pulse'　　مريض mæríz 'sick'

14.23 z has the independent shape: ض

رياض riáz 'Riyadh'　　قرض qærz 'debt'

مرض mæráz 'disease'　　مقروض mæqrúz 'indebted'

15. t and z (ta and za) have the same basic form. The ta is dotless; the za has one dot above.

15.10 t- has the initial shape: ط

طبيب tæbíb 'physician'　　طلا telá 'gold'

طوطی tutí 'parrot'　　اطاق 'otáq 'room'

15.11 -t- has the medial shape: ـط

سطل sætl 'pail'　　مطر metr 'meter'

نقطه noqté 'point'　　ارسطو 'æræstú 'Aristotle'

15.12 -t has the final shape: ـط

ربط ræbt 'connection'　　وسط væsét 'middle'

خط xætt 'line'　　رابط rabét 'liaison'

15.13 t has the independent shape: ط

احتياط 'ehtiyát 'caution'　　اختلاط 'extelát 'mixture'

ارتباط 'ertebát 'communication'　　سقراط soqrát 'Socrates'

15.20 z- has the initial shape: ظ

ظرف zærf 'vase'　　ظهر zohr 'noon'

ظاهر zahér 'apparent'　　وظيفه væzifé 'duty'

15.21 -z- has the medial shape: ـظ

نظارت nezarét 'supervision'　　نظر næzér 'view'

انتظار 'entezár 'expectation'　　منظور mænzúr 'purpose'

15.22 -z has the final shape: ظ

خدا حافظ. xodahaféz 'Good-bye!'	حفظ hefz 'protection'
تلفظ tælæffóz 'pronunciation'	حافظ haféz 'Hafez'

15.23 z has the independent shape: ظ

محظوظ mæhzúz 'pleased'	لحاظ leház 'point of view'
ملفوظ mælfúz 'pronounced'	الفاظ 'ælfáz 'words'

16 The following basic shape ('eyn) represents ' (that is, either a glottal stop or stricture, or a lengthening of the preceding vowel). Initially it functions much like 'æléf, but occurs only in words of Arabic origin.

'eyn and qeyn are alike in their basic form, but qeyn has one dot above.

16.10 '- has the initial shape: ع

عاقل 'aqél 'wise'	عبارت 'ebarét 'phrase'
عجیب 'æjíb 'strange'	اجتماعی 'ejtema'í 'social'

16.11 -'- has the medial shape: ـعـ

سعی sæ'y 'attempt'	بعد bæ'd 'after'
فعل fe'l 'verb'	معروف mæ'rúf 'famous'

16.12 -' has the final shape: ع

ربع rob' 'quarter'	قلع qæl' 'tin'
جمع jæm' 'plural'	مایع mayé' 'liquid'

16.13 ' has the independent shape: ع

مضارع mozaré' 'present (tense)'	اجتماع 'ejtemá' 'gathering; society'
دفاع defá' 'defense'	شجاع šojá' 'brave'

16.20 q- has the initial shape: غ

غاز qaz 'goose'	غذا qæzá 'food'
روغن rowqén 'oil, fat'	کاغذ kaqéz 'paper'

16.21 -q- has the medial shape: ـغـ

مغز mæqz 'brain, marrow'	بغل bæqél 'armpit'
تغییر tæqyír 'change'	مشغول mæšqúl 'busy'

16.22 -q has the final shape: غ

میغ miq 'cloud, mist' (literary)	تیغ tiq 'blade'

16.23 q has the independent shape: غ

دروغ dorúq 'lie, false'	دماغ dæmáq 'nose'
یوغ yuq 'yoke'	استفراغ 'estefráq 'vomiting'

LIGATURES

17. Ligatures are a combination of two letters into a specialized form which is not a part of the regular Persian alphabet. They are limited to the following combinations: lam-'æléf, kaf-'æléf, gaf-'æléf, kaf-lám, and gaf-lám. When the sequences kaf-lam-'æléf and gaf-lam-'æléf appear, only two of the letters may be combined into a ligature. The remaining letter is printed in its usual form.

17.10 The combination lam-'æléf, when in isolated position, is printed: لا

لاستیک lastík 'rubber' لاک lak 'sealing wax'

لاشخور lašxór 'vulture' اولاد 'owlád 'child'

17.11 When in final position, the combination lam-'æléf has this shape: ـلا

خلال xelál 'interval' (literary) اختلاف 'extelâf 'difference'

فلاحت fælahét 'agriculture' طلا telá 'gold'

17.20 The combinations kaf-'æléf and gaf-'æléf, when in isolated position, are printed ک and گ. When in final position, they are printed ـک and ـگ respectively.

کاغذ kaqéz 'paper' انکار 'enkár 'denial'

گام gam 'step' نگاه negáh 'glance'

17.30 The combinations kaf-lám and gaf-lám are written ک and گ initially, ک and گ medially, کل and گل finally, and کل and گل when independent.

کلم kælém 'cabbage' گلو gælú 'throat'

الکلی 'ælkolí 'alcoholic' انگل 'æŋgél 'parasite'

شکل šekl 'shape' انگلیسی 'engelisí 'English'

توکل tævækkól 'trust (e.g., in God)' گل gol 'flower'

17.40 The combinations kaf-lam-'æléf and gaf-lam-'æléf are more commonly printed کلا، گلا؛ than کلا، گلا.

کلاغ = کاغ کلاغ kæláq 'crow' بکلاهش = بکاهش bekoláhæš 'to his hat'

گلابی = گابی golabí 'pear' بگلابی = بگابی begolabí 'to the pear'

SPECIAL SIGNS

In Sections 3 and 4 some of the means for indicating vowels and diphthongs were mentioned. In the following nine sections, examples and further combinations will be discussed.

18. The vowels æ, e, and o may be represented by the following signs:

æ ـَـ (placed above the preceding consonant):

تَر tær 'soak' نَر nær 'male'

e ـِـ (placed under the preceding consonant):

کِرم kerm 'worm' شِکل šekl 'shape'

o ـُـ (placed above the preceding consonant):

پُر por 'full' تُرُب torób 'radish'

Absence of vowel may be indicated by the sign ـْـ placed above the consonant to which it belongs:

پِدَرْ pedær 'father' اَگَرْ 'ægær 'if'

These four signs are, however, hardly ever used in printed (or written) Persian. These signs are sometimes called 'vowel points'.

19. 'ælef mæddí. The form آ, called 'ælef mæddí (essentially an 'ælíf drawn horizontally above a conventional 'ælíf) expresses the sound a when in independent or final position. In alphabetical listing, 'ælef mæddí counts as 'ælíf plus 'ælíf.

19.10 'ælef mæddí in independent position: آ

آب 'ab 'water' آرد 'ard 'flour'

آزاد 'azád 'free' آسیا 'asiyá 'Asia'

19.11 'ælef mæddí in final position: آ

مآثر ma'asér 'memorable deeds' (literary) تآتر teátr 'theater'

20. Initial 'ælíf plus vav expresses 'u, 'ow, or 'v: او

اورنگ 'owræng 'throne' (literary) او 'u 'he, she'

اولاد 'owlád 'children' اول 'ævvæl 'first'

21. Initial 'ælíf plus ye expresses 'i, 'y, or (with a vowel of course) 'ey or 'æy: ای

ایران 'irán 'Iran' ایوان 'eyván 'palace'

اینجا 'injá 'here' ایشان 'išán 'they'

22. hæmzé. The symbol ٔ, called hæmzé, is sometimes used over ا, و, ی, and ه, as described in the following sections:

22.10 A glottal stop in the middle of a word is sometimes indicated by the small sign ٔ (hæmzé) above و, ا, ی (without dots), and ه.

سائل sa'él 'beggar' رؤسا ro'æsá 'chiefs'

مسئول mæs'úl 'responsible' امریکائی 'emrika'í 'American'

There is a tendency to omit the hæmzé and use a ی (ye) instead everywhere except in words of Arabic origin.

22.20 'æleʃ hæmzé. There are also cases where, in the middle of a word, hæmzé is written above an 'æléf (though its omission in print is frequent) to express the sequence æ'. Note that 'æléf here does not express the sound a.

تأخیر tæ'xír 'delay' تأدیب tæ'díb 'chastisement'

تأنی tæ'nní 'slowness, hesitation' مأخذ mæ'xéz 'source, origin'

22.30 hæmzé is also used over a final h (ۀ) (though omission in printing is frequent) to indicate connecting e ('ezafé) in the case of final written but unpronounced h (see Section 3). The hæmzé in this case has the pronunciation ye.

خانهٔ بزرگ xanéye bozórg 'the big house'

دستهٔ کاغذ dæstéye kaqéz 'the bunch of paper'

Final connected ه may have ٔ (hæmzé) to represent indefinite -i: هٔ 'i. This is increasingly being written ای 'i.

خانهٔ xané'i 'a house' دستهٔ dæsté'i 'a bunch'

23. tæšdíd. The doubling of consonants, in the middle or at the end of a word, may be indicated by the small sign ّ (tæšdíd) written above the letter to be doubled; but it, too, is often omitted.

اوّل 'ævvæl 'first' مخفّف moxæffæf 'contraction'

معلّم mo'ællém 'teacher' تلفّظ tælæffóz 'pronunciation'

24. tænvín. An ending -æn (of Arabic origin) of certain adverbs is indicated by the following sign ً (tænvín), written usually above a final 'æléf. Like hæmzé and tæšdíd, it, too, is often omitted. Note that in this case, final 'æléf does not have the value a.

مثلاً mæsæláen 'for example' اولاً 'ævvæláen 'firstly'

تقریباً tæqribáen 'almost, approximately' قبلاً qæbláen 'beforehand, earlier, before'

25. In older printed material, the final and independent h, ه and ه, appear, at the end of some words, with two dots above them. These words are invariably of Arabic origin. In modern writings these forms are very rarely found. In words where the dotted h is pronounced t, the symbol ة or ت is used.

اصالت 'esaláet 'genuineness' زیارت ziyaráet 'pilgrimage'

In words where an originally dotted h is retained without dots, as ه, it represents the vowel e:

اجازه 'ejazé 'permission'

26. Mention was made in Section 6.35 of the fact that final y is sometimes pronounced as if it were final 'æléf, and is pronounced a.

حتی	hætta 'even though'	اعلی	'æ'lá 'higher, highest'
ماوی	mæ'vá 'dwelling, abode'	سکنی	sokná 'dwelling'

But note:

مبنی / مبنا	mæbná 'basis, foundation'	مبنی	mæbní 'based, founded'

However, when these words are followed by the syllable -ye (another form of -e, the 'ezafé sign), the letter ye is replaced by 'æléf. This 'æléf represents the sound a (which is the same sound as was indicated by the ye), and the letter ye indicates (as it does in other instances) the y of the 'ezafé ending -ye.

اعلی	'æ'lá	اعلای	'æ'láye
ماوی	mæ'vá	ماوای	mæ'váye
سکنی	sokná	سکنای	soknáye

NUMBERS

27. In Persian the numerals are written and printed from left to right. They have the following shapes:

یك	yek 'one'	۱	شش	šeš 'six'	٦
دو	do 'two'	۲	هفت	hæft 'seven'	۷
سه	se 'three'	۳	هشت	hæšt 'eight'	۸
چهار	cæhár 'four'	٤	نه	noh 'nine'	۹
پنج	pænj 'five'	٥	صفر	sefr 'zero'	۰

What has been described in the preceding pages is the Persian writing system as it appears in newspapers, periodicals, books, and other material printed on the present-day presses of Iran. In order to acquaint the reader with typewritten Persian, and to give him a general idea about handwritten Persian, a short note on each follows.

The statements made in the notes on combination of words and on punctuation apply to all three forms of the Persian writing system described in the present paper.

TYPEWRITTEN PERSIAN

0. The main difference between printed Persian and typewritten Persian consists in the number of the forms of each letter.

1. All the non-connectors, that is, 'æléf, dal, zal, re, ze, že, and vav, each of which has two forms in printed Persian (Section 4), and two of the connectors, ta and za, each of which has four printed forms (Section 15), have only one form each on the typewriter keyboard:

ا د ذ ر ز ژ ط ظ

2. The letters be, pe, te, se, nun (Section 6); jim, ce, he (hottí), xe (Section 7); sin, šin (Section 8); sad, zad (Section 14); fe, qaf (Section 10); kaf, gaf (Section 9); lam (Section 11); and mim (Section 12), which have four forms each in printing, have only two forms each on the typewriter, an initial-medial form and a final-independent form:

Initial-medial:

b	p	t	s	n	j	c	h	x	s	š	s	z	f	q	k	g
بـ	پـ	تـ	ثـ	نـ	جـ	چـ	حـ	خـ	سـ	شـ	صـ	ضـ	فـ	قـ	کـ	گـ

Final-independent:

b	p	t	s	n	j	c	h	x	s	š	s	z	f	q	k	g
ب	پ	ت	ث	ن	ج	چ	ح	خ	س	ش	ص	ض	ف	ق	ک	گ

3. The letter he (hævváz) (Section 13), which is a connector, has three forms in typewritten Persian: initial, هـ; medial, ـه; and independent-final, ه.

4. The letter ye (Section 6.30) has three forms: initial-medial, یـ; final, ی; and independent, ﻯ.

5. Two letters, 'eyn and qeyn (Section 16) have four forms each (just as they do in printed Persian): initial, عـ; medial, ـع; final, ع; and independent, ﻉ.

6. Of the ligatures (Section 17), the typewriter has only lam-'æléf: لا ﻻ.

7. There is a hæmzé (Section 22) which can be placed above 'æléf, vav, or ye. There is also an initial-medial ء.

8. The_mæddí sign (Section 19), which in printed Persian is part of the type آ, is a separate (dead) key on the typewriter.

9. There is a dotted he hævvæz on the typewriter.

HANDWRITTEN PERSIAN

0. As is natural, Persian letters, which have basically the same form in printing, typing, and handwriting, are somewhat different and less easily recognizable in handwriting. This is partly due to the differences among individual writers, but there are also generally accepted written forms of the letters and combinations of letters which are quite different from the printed shapes.

What follows is a description of some of the major differences between a variety of handwritten Persian, which is more or less typical, on the one hand, and printed (and typewritten) Persian, on the other. No attempt will be made to discuss calligraphy or any of a number of other types of handwritten Persian.

1. The three pointed strokes of sin and šin can, in all their forms, be replaced by a slightly curved straight line:

سر sær 'head' بشر bæšær 'mankind'

However, when there are two successive sins, or šins, or one of each, it is preferred that only one of them be written in the above manner.

2. The final curved stroke of the final and independent sin, šin, sad, zad, and ye may be replaced by a stroke slanting downward to the left and ending in an upward turn to the left. The entire stroke may be angular or slightly curved:

یاس yas 'jasmine' کیش kiš 'religion'

خاص xas 'particular' مریض mæríz 'ill'

روی rúye 'on' بی bi 'without'

3. The curved stroke of the final or independent nun may be replaced by a stroke slanting downward to the left, as in 2 above, but ending in a turn to the right:

آن 'an 'that' من mæn 'I'

4. A preceding connector joins jim, ce, he (hottí), and xe from the top, not from the side, as in typewritten and most printed forms:

بچه bæccé 'child' نخ næx 'thread'

5. A medial be, pe, te, se, nun, and ye before final nun, and sometimes before re, ze, and že, has the shape ـ, coming up from the preceding letter and joining the next by a short straight line:

بین beyn 'between' پیر pir 'old'

6. When followed by final re, ze, že, or one of the final forms described in 1, 2, and 3 above, the pointed stroke of a medial be, pe, te, se, nun, and ye is replaced by a curve:

پیر pir 'old' نبی næbí 'prophet'

شیر šir 'milk' چین cin 'China'

7. Care must be taken not to form a curve such as described in 6 above when joining an initial to a final letter, since the curve may be taken for a letter itself. Compare, for example,

پیر pir 'old' and پر por 'full'

In joining an initial letter to one of the final forms described in 1, 2, and 3 above, the angle between the letters should be rather sharp. Otherwise, any resulting curve may be taken for a letter. Compare, for example,

کش koš 'Kill!' and کشتی kəští 'ship'

8. Initial he (hævvéz) may be written like a mim with a comma-like addition below it:

هنر (for هنر) honér 'art'

This 'comma' stroke is written from the bottom up: ۷

9. Final he (hævvéz) is a simple curve downward, the central part of which is only slightly (if at all) raised:

چه če 'What?'

10. If necessary, any medial letter and, less frequently, an initial letter, may be stretched, usually to fill in the line:

کتاب ketáb 'book' ظاهر zahér 'apparent'

It is not very common to stretch two successive letters in the same word; nor is it considered permissible to stretch a letter when it is likely to cause ambiguity, as in the case of پر por 'full', where, if the pe is stretched, the stretching line may be taken for a medial sin (see 1 above), thus resulting in the word پسر pesér 'boy'.

11. All forms of kaf have one oblique bar. This bar is added after the word is written, like the crossing of a t in English. (See also 14 below.) The ' which appears inside the final and independent forms of printed kaf and gaf does not appear in handwriting.

12. The ligature lam-'æléf is written ـلا final, and لا independently:

طلا telá 'gold' بالا balá 'above'

Sometimes lam-'æléf is written ـلـا or لـا :

طلا telá 'gold' بالا balá 'above'

13. The following observations may be made concerning dots:

(a) Two dots belonging to the same letter can be replaced by a straight line:

یا ya 'or' تا ta 'to, till'

(b) Three dots belonging to the same letter can be replaced by a line with a dot above or below it (depending on whether the dots are placed above or below the letter):

پدر pedér 'father' ثابت sabét 'fixed'

Three dots may also be replaced by a small circle, thus: پدر, ثابت.

(c) The dots are frequently not placed exactly above or below the letter to which they belong:

افتخار 'eftexár 'pride, honor' بخاری boxarí 'stove, heater'

In cases where there is ambiguity, the context usually helps. Compare

بخاری boxarí 'stove' and نجاری næjjarí 'carpentry'

both of which can be written نجاری.

(d) Not infrequently, dots are left out altogether. In such cases, the context usually helps in the reading.

14. The dots, hæmzés, and all the special signs are added after the whole word, or any part of a word ending in a non-connector, is written.

15. The numerals ٤ and ٦ are written ۴ and ۶.

COMBINATION OF WORDS

0. Compound words are ordinarily, but not necessarily, written as one word (unless the first word ends in a non-connector). Certain other words are treated as compounds in writing, that is, they are joined together. A few of the commoner cases are listed below:

1. این 'this' is sometimes joined to a following short word:

اینگونه اینکتاب ایندرس اینوقت

This does not happen when the following word begins with a letter of the group ب پ ت ث ز ژ ذ.

The above statement is true, though much less frequently, of آن 'that':

آنکتاب آنوقت آنگه آنروز

2. What was said above of این is true of یك 'one', too.

3. The combinations باین ,بان and بدین ,بدان may join a following short word:

باینوسیله بانوسیله بدینوسیله بدانوسیله

4. If این follows a preposition ending in a consonant, it may lose its 'eléf, in which case the two words will be combined:

ازین (for از این) برین (for بر این)

This is true of آن, too; only here it is the mæddí (which, the reader will remember, was an 'eléf originally) which is omitted:

ازان (for از آن) بران (for بر آن)

5. Some writers drop one of the two h's (he hævvéz) when they add the plural suffix to a word ending in an h (he hævvéz) that represents the vowel e (see Section 13.12):

خانها (for خانه‌ها)

However, even these writers preserve both the h's if dropping one of them may cause confusion, as in the case of

نامه‌ها 'names'	and	نامها 'letters'
نشانها 'medals'	and	نشانه‌ها 'signs'
خانها 'khans'	and	خانه‌ها 'houses'

6. When می joins a verb form beginning with an ا ('æléf) or a ی (ye), the 'æléf or ye of the verb root may be dropped:

میفتد (for می‌افتد) میابد (for می‌یابد)

The dropping of 'æléf may be considered as indicating the fact that the glottal stop represented by ا ('æléf) is omitted, though even when the 'æléf is preserved, it may quite often be ignored in pronunciation.

7. Some of the cardinal numerals one to ten may join certain words following them (mostly the words سال 'year' and ماه 'month'):

یکسال (for یك سال) 'one year'

پنجساله (for پنج ساله) 'five-year-old'

ششماهه (for شش ماهه) 'six-month-old'

These numerals may or may not join to صد '100', هزار '1,000', and, occasionally, میلیون '1,000,000'.

8. The word است has two written forms: است and ست (or ـست). Used after a consonant, the latter represents æst combined with the preceding word:

خوبست. xúbæst 'It is good.'

After a vowel, ست represents st:

کتابست. ketábist 'It is a book.'

However, even if the ا ('æléf) is retained in such cases, it usually has no sound value; thus, کتابی است is pronounced exactly like کتابست.

PUNCTUATION

0. The following punctuation marks are used in Persian.

.	period
؟	question mark
!	exclamation point
:	colon
؛	semi-colon
،	comma
—	dash
« »	quotation marks
( )	parentheses
[ ]	brackets
ـ	hyphen

These marks are not consistently used in Persian. There is variation among different writers, and inconsistency may occur in the work of a single writer. When they are used, they generally follow Western, often French, usage. The following notes are meant as a guide, not a description of all possible uses.

1. The PERIOD (.) is used at the end of a declarative or imperative sentence, and sometimes after an interrogative one:

برادرم را دیدم. 'I saw my brother.'

بنشینید. 'Sit down.'

کجا میروید. 'Where are you going?'

Several periods indicate an omission:

آقای . . . اینجا نیست. 'Mr. ... is not here.'

2. The QUESTION MARK (؟) is used, as in English, after questions:

اسمتان چیست؟ 'What is your name?'

It may also be used after an indirect quotation which is a question itself or of which the original was a question:

ازو پرسیدم مرا میشناسد یا نه؟ 'I asked him whether he knew me or not.'

3. The EXCLAMATION POINT (!) is used, as in English, after sentences expressing surprise, astonishment, emphasis, etc.:

این اصفهانی، اصفهانی نیست! 'This Isphahani is not an Isphahani!'

فراموش نکنی! 'Be sure not to forget!'

It is also used after imperatives:

برخیز! 'Stand up!'

4.. The COLON (:) is used

(a) before a list, when this list is at the end of the sentence:

این کتابهارا خریدم:. . . . 'I bought these books:...'

(b) before a quotation, direct or indirect:

دوستم گفت: «کجا میروید؟»
دوستم گفت: کجا میروید؟ } 'My friend said: "Where are you going?"'

دوستم از من پرسید: کجا میروید؟ 'My friend asked me where I was going.'

(Note that quotation marks are optional.)

(c) before an explanatory note referring to the preceding sentence:

همه جارا جستجو کردم، ولی اورا پیدا نکردم: از شهر خارج شده بود.

'I looked everywhere but I didn't find him: he had left the city.'

5. The SEMI-COLON (؛) may be used instead of the colon before an explanatory note (see 4c above):

همه جارا جستجو کردم، ولی اورا پیدا نکردم؛ ازینجا رفته بود.

'I looked everywhere but I didn't find him; he had left here.'

6. The COMMA (،) is used to separate phrases, clauses, or words in a list (usually equivalent to single bar juncture, |):

کتابها، مدادها، قلمها 'the books, the pencils, the pens'

بخانه رفتیم، شام خوردیم، بعد خوابیدیم.

'We went home, had supper, and went to bed.'

It may indicate a parenthetical phrase or clause:

احمد، برادرم، اینجاست. 'Ahmad, my brother, is here.'

Sometimes this indication of the phrasing (juncture) is necessary to avoid ambiguity, as:

این تخته سیاهست. 'This board is black.'

این، تخته‌سیاهست. 'This is the blackboard.'

In cases like the following, the comma is used to avoid ambiguity caused by two successive identical words:

این اصفهانی، اصفهانی نیست! 'This Isphahani is not an Isphahani!'

7. The DASH (—) is used like the comma (see 6 above). It is also used in dialogue to mark the beginning of a speaker's remarks:

کتابها— قلمها— مدادها 'the books, the pens, the pencils'

این تخته سیاهست. 'This board is black.'

این— تختهسیاهست. 'This is the blackboard.'

این اصفهانی—اصفهانی نیست! 'This Isphahani is not an Isphahani!'

احمد— برادرم— اینجاست. 'Ahmad—my brother—is here.'

— شما کجایی هستید؟ 'Where are you from?'

— فرانسوی. 'France.'

A dash may also be used to indicate an omission, instead of several periods (see 1 above):

آقای — اینجا نیست. 'Mr.——is not here.'

8. QUOTATION MARKS (« ») may be used
(a) to mark a direct quotation:

پدرم گفت: «کتابرا بمن بدهید.» 'My father said, "Give me the book, please."'

(b) to indicate a foreign word, or for emphasis (as in English):

میدانید «chair» یعنی چه؟ 'Do you know what "chair" means?'

(c) to set off a parenthetical phrase or sentence:

احمد «برادرم» شام میخورد. 'Ahmad, my brother, is having his supper.'

9. PARENTHESES (( )) may be used instead of quotation marks:

پدرم گفت (کتابرا بمن بدهید.) 'My father said, "Give me the book, please."'

میدانید (chair) یعنی چه؟ 'Do you know what "chair" means?'

احمد (برادرم) شام میخورد. 'Ahmad, my brother, is having his supper.'

10. BRACKETS ([ ]) are generally restricted to scholarly works, where they are used as in English:

«شکسپیر از مولفین بزرگ قرن چهاردهم [کذا فی‌الاصل] بود.»

'Shakespeare was one of the great writers of the 14th [sic] century.'

11. The HYPHEN (-) is not commonly used, but occurs occasionally to indicate the division of a word at the end of a line:

اگر از رمان خوشتان میاید میتوانید این کتاب— 'If you like novels, you can take
هارا با خود بمنزل ببرید. these books home with you.'

However, words are almost never divided at the end of a line in handwriting and rarely in print. When such a division does occur, it is usually done between a root and an affix, or between affixes, providing the break occurs after a non-connector (see Section 4) and the affix to begin the next line starts with a consonant:

در سالهای اخیر خیلی از شعر های اورا خوانده‌ام. 'I have read many of his poems in recent years.'

You may not, for example, divide خریدم into خرید xærid- and م -æm, as the latter begins with a vowel.

Compound words, the members of which occur also as separate words, may be divided between the two parts of the compound:

مهمانخانه = مهمان‌خانه 'hotel'